Modelling Objectively

Nancy Cartwright (Durham University and University of California, San Diego) and Faron Ray (University of California, San Diego)

CHESS Working Paper No. 2023-01
Durham University
February 2023

CENTRE FOR HUMANITIES
CHESS
ENGAGING SCIENCE AND SOCIETY

CHESS working paper (Online) ISSN 2053-2660

Modelling Objectively[1]

Nancy Cartwright (Durham University and University of California, San Diego) and Faron Ray (University of California, San Diego)

Nancy Cartwright
Centre for Humanities Engaging Science and Society (CHESS), Department of Philosophy, Durham University,
50 Old Elvet, Durham, DH1 3HN, UK
nancy.cartwright@durham.ac.uk

Faron Ray
University of California, San Diego, Ridge Walk Academic Complex - Arts & Humanities Bldg. Room 0434, UC San Diego
9500 Gilman Dr. La Jolla, CA 92093, USA
fray@ucsd.edu

[1] This research was funded by the John Templeton Foundation as part of The Successes and Failures of Science Through the Lens of Intellectual Humility: Perspectives from the History and Philosophy of Science.

Abstract

When models are used to guide decisions that can affect people's lives, those involved in modelling have a duty of care towards those who are likely to be affected by those decisions. Drawing on Cartwright et al's (forthcoming) account of objectivity we argue that this duty of care should be understood as involving a duty to *model objectively*. We illustrate with two tragic cases for which there was a conspicuous *failure* to model objectively. As these cases bring out, however, objective modelling does not occur in a vacuum. Modelling takes place embedded within a tangle of institutional norms, expectations, habits, patterns and practices, and modelling objectively will often be a *systems-level problem* in need of *systems-level solutions*. We end by diagnosing one such potential systems-level problem—a failure of *intellectual humility* in the institutions in which modelling takes place.

1. Introduction

When models are to be used to guide decision-making about real world happenings, this creates a *duty of care* towards those whose lives the decision-making is likely to affect. In this paper, we argue that this duty should be understood as involving a duty to *model objectively.* Borrowing Cartwright et al's (forthcoming) account of 'Objectivity to Be Found', we draw a close connection between objectivity and *getting it right,* where getting it right demands both i) finding the right purposes, and ii) finding the right methods given those purposes.

Our approach differs from other recent discussions of duties in modelling which focus on a modeller's duty to *clarify* the intended purposes of a model (Winsberg and Harvard 2022). Whilst clarifying purposes is important, we think that the duty of care demands more. We illustrate with an instance of modelling for the Vajont Dam in the Dolomites, the site of the 1963 disaster that killed over 2000. Here the engineers carrying out the modelling failed their duty of care and yet there was no lack of clarity as to the intended purposes of the modelling. Instead, the problem was that the engineers and the institutions overseeing the design and building of the dam focused on the *wrong* purposes; purposes that emphasised ensuring that the dam would hold in spite of subsidy and landslides. Ensuring that the inhabitants of the towns and villages around the dam would be safe *if* a landslide were to occur unfortunately received much less attention.

We close by identifying a problem that we think makes failures to model objectively all too natural: a failure of institutional humility in the institutions in which the modelling is embedded, e.g. the institutions that train modellers or hire them or regulate or otherwise affect modelling practices—an inappropriate *presumption* that their practices are the right practices.

The paper is structured as follows. §2 introduces the duty of care in modelling. §3 introduces Cartwright et al's account of 'Objectivity To Be Found'. §4 uses the account to spell out a modeller's duty of care and highlights the important differences between our proposed focus on *finding* the right purposes and a focus on clarifying intended purposes. §5 illustrates with a case of unobjective modelling—two instances of modelling conducted for the Vajont Dam. §6 turns to the question of why failures to model objectively are often so natural in certain circumstances, where we conjecture that this is at least partly explained by failures of intellectual humility in the institutions in which the modelling is embedded. §7 concludes.

2. A Duty of Care in Modelling

On 15th April 1989, 96 supporters were killed and 766 injured at Hillsborough Football Stadium in Sheffield, South Yorkshire in the UK when a fatal crush occurred in the stadium's enclosed pens. South Yorkshire Police[2] were *defacto* in charge of crowd safety and thus had both a moral and legal *duty of care* to ensure that those entering the ground were not exposed to unreasonable levels of risk.[3] They failed to fulfil this duty. What went wrong? One could explain the failure by tracing the chain of mistakes made by various officers on the day. First, officers stationed outside the ground lost control of crowds as they waited to enter, leading to the onset of a dangerous crush by Leppings Lane turnstiles. Next, to ease the crush outside a senior officer requested for an exit-gate (Gate C) to be opened in order to allow maximum flow into the ground. Finally, Chief Superintendent David Duckenfield granted the officer's request, leading crowds of supporters to be funneled down the natural channel that existed from Gate C to one of the already full enclosed pens. Whilst accurate, however, this story only goes so far as an explanation for South Yorkshire Police's failure to fulfil their duty of care, for the failures of the South Yorkshire Police as an institution were not limited to the actions taken by individual officers that day. Rather, they extended to the manner in which South Yorkshire Police set about *preparing* to act. That is, they extended to the *modelling* conducted by South Yorkshire Police.

Talking about the failures of South Yorkshire Police in terms of modelling may seem odd. There were no mathematical equations to solve, no computerised simulations being run; the information gathered was almost exclusively stored in the lead officers' heads. Yet, to understand this failure to fulfil the duty of care one must nevertheless look to the pre-match preparation conducted by South Yorkshire Police and this *was* a case of modelling.

Preparation for the FA Cup semi-final at Hillsborough that day included no attempt to estimate the amount of people that could be safely let into the stadium's pens. Nor did it involve any attempt to gather vital information about patterns of crowd flow *into* the pens, information that would have been available from past match reports and from the past experience of previous lead officers at the ground.[4] Finally, pre-match preparation did not involve using this would-be information to estimate how crowd flow patterns would change under various safety-critical

[2] 'South Yorkshire Police' here is meant loosely including not only the police officers themselves but the institutions charged with hiring, training and overseeing them.
[3] The duty of a care is a foundational concept in Tort Law closely tied to the notion of *fault liability.* The most famous articulation of the duty of care is that given in Donoghue v Stevenson ([1932] AC 562): "You must take reasonable care to avoid acts or omissions which you can reasonably foresee would be likely to injure your neighbour […] persons who are so closely and directly affected by my act that I ought reasonably to have them in contemplation […]" See: https://www.bailii.org/uk/cases/UKHL/1932/100.html.
[4] There had been severe overcrowding (injuring 38) seven years prior to the disaster. See Taylor (1989, 21).

interventions, e.g., the opening of the particular exit-gate which would eventually be opened with fatal consequences.[5]

Whilst the failures of South Yorkshire Police at Hillsborough might seem a far cry from issues of scientific modelling, the principles are much the same. Officers' actions that day were guided by models just as policy interventions or engineering projects are guided by models. Indeed, acting intelligently in any domain will require that you have a model, or models, of that domain. Now, since South Yorkshire Police had a duty of care to those in the stadium that day, so too did this duty of care extend to their modelling. Indeed, it was precisely that the police were working with *bad models*[6] that led to the catastrophic mistakes they made that day.

So, what exactly was so bad about the modelling carried out by South Yorkshire Police? They had not estimated certain factors, sure, but then they had not estimated lots of things, many of which turned out to be inconsequential. What was it about the pre-match modelling by South Yorkshire Police that proved so fatal that day? And what practices might have ensured that South Yorkshire Police had modelled 'correctly 'rather than 'incorrectly'? The sections to follow will aim to sketch general answers to these kinds of questions. In particular, we will spell out some of what we believe the duty of care requires in the context of modelling. In doing so we will make heavy use of the concept of *objectivity.* Specifically, we will argue that a duty of care requires that modellers *model objectively.*

Before we move to discuss what it means to model objectively, we must first say a few words about the concept of objectivity itself. We begin below by addressing some of the recent criticism that has been levelled *against* the use of 'objectivity' in science before moving to say how we suggest the concept should be understood.

3. Why Objectivity?

Any list of philosophically fraught concepts is likely to include 'objectivity' somewhere amongst the ranks. Ian Hacking puts it bluntly, claiming that 'objectivity' should be thrown in the bucket alongside 'reality' and 'rationality' as "fancy words conceived in philosophical sin" (Hacking 2015, 4). Worries about 'objectivity' come in two main forms. The first is that the concept is too sprawling, too full of content, too *unwieldy* to play any useful role in science. The second is that the concept is too bound up with the undesirable notion of *value freedom* to do useful work. According to this second worry, 'objectivity' has unhelpful baggage and as such is more trouble than it is worth. Since we believe that the concept *can* do useful work in science and shall be putting the concept to such work in the discussion to follow, it will pay to say a few words about these complaints before introducing the account of objectivity that we endorse.

[5] During the 2015 inquests, Chief Superintendent Duckenfield claimed that the biggest mistake he made on the day was not 'foreseeing' where fans would go when he opened Gate C, calling this failure of foresight "one of the biggest regrets of his life" (see Atkinson 2016). Whilst Duckenfield reports that he regretted this lack of foresight, the fact that he lacked such foresight ultimately stems from failures to investigate likely crowd flow patterns in advance of the match.

[6] Perhaps more accurately 'an incomplete set of models'.

The worry that objectivity is too sprawling and ill-defined to play any useful role in science has been voiced most clearly by Hacking (2015). According to Hacking, 'objectivity' is simply a catch-all for a tangle of largely unrelated epistemic virtues. To say that X is or is not objective, according to Hacking, is simply to say that X possesses or lacks this or that epistemic virtue. 'Objectivity', therefore, can always be *replaced* by something more specific. Better get down to specifics, concludes Hacking, than trade in redundant talk of 'objectivity'.

Hacking's complaint gets something right. To the degree that praising a colleague's work for its objectivity and reprimanding another's for its *lack* of objectivity is shorthand for more substantive praise or criticism, it is undeniable that the concept has little, if any, value. Yet, such descriptive applications of the adjective 'objective' fall far short of exhausting the ways in which the objectivity concept can be used. Consider the injunction, 'be objective'. Is reference to objectivity here *replaceable* by something more specific? No, far from it. Rather, just like 'respectful' in a parent's injunction to their child to *be respectful,* the prescriptive use of objectivity does not need to refer to a single activity or set of activities in order to play a useful role. Indeed, we take it that it is the very open-endedness of injunctions to *be objective* just like those to *be respectful* that give them their action-guiding role; it is part of the injunction to be objective that one should *find out* what it takes to be objective in each individual case. Much more will be said about this notion of *finding* below.

What about the second worry, namely, that 'objectivity' is too bound up with the undesirable notion of *value freedom*? According to Brown (2019), such connotations ultimately make demands for objectivity counter-productive; whilst philosophers of science are keen to point that science is rife with ethical decision points,[7] demands for objectivity imply the mechanical following of procedure and the *removal* of all value judgements.[8] Accordingly, whilst demands for researchers to 'be objective' are not strictly *redundant* as Hacking worries, it is nonetheless the case that the concept is more trouble than it is worth. Indeed, its being wrapped up with notions of mechanical rule-following and a removal of value judgements makes the concept positively *harmful*—it encourages scientific ideals that are at odds with the way (good) science is done.

This worry is well motivated, yet we think that it overestimates the counterproductive baggage of objectivity-talk. Far from objectivity implying an absence of value judgments and a blind commitment to procedure, we suggest that in many cases complaints that someone is not *being objective* will, in fact, constitute complaints that one has not used their judgement *properly*—'Wait, you just blindly follow procedure? That's not being objective! You must use your considered judgement!' In fact, what everyday injunctions to be objective demand, we take it, is not that one *avoids* the influences of judgement but that one uses their judgement *correctly* given the demands of the case. Far from being in opposition to the presence of value judgments, injunctions for

[7] As Richard Rudner (1953, 2) pointed out, something as apparently mundane as setting the criteria for acceptance or rejection of a hypothesis will itself "be a function of the importance, in the typically ethical sense, of making a mistake in accepting or rejecting the hypothesis." See also Douglas (2000).

[8] The concept of objectivity has, of course, a rich history and the baggage associated with objectivity here is most associated with what Daston and Galison (2007) refer to as *mechanical objectivity*—a notion that has its origins in the 19th century.

researchers to be objective are very often injunctions that those researchers make *good calls* and good calls imply good value judgements. What 'good' means here will come out below.

So, how do we suggest 'objectivity' be understood? First, we think it is important to recognise that the fuzziness of the concept is here to stay. The concept of objectivity as it shows up in science, law and everyday natural language is what Otto Neurath would call a *Ballung concept*;[9] it is sprawling, dense and unruly, with many separately identifiable senses.[10] Following the suggestion of Eleonora Montuschi (further developed in Cartwright et al (forthcoming)) we call this loose sense of objectivity 'Objectivity As We Know It'. As noted above, however, uses of 'objectivity', like uses of other Ballung notions, needn't be otiose. Used prescriptively, objectivity can do real work. For another analogy, it is common in commercial contracts to stipulate that parties use their 'best endeavours' to bring about what has agreed to be done and companies and individuals can be on the hook for paying significant damages if they can be shown *not* to have used their 'best endeavours'. But what *are* such 'best endeavours'? Whilst this will be a matter fixed by context, this by no means implies that the concept lacks usefulness. Similarly, we suggest, for prescriptions to 'be objective'. Indeed, we suggest it is instructive to draw a tight connection between objectivity and *getting it right.* Given the circumstances, there will often be a fact of the matter as to whether one has *got it right*, but there is no clear notion of 'getting it right' that can be defined free from context. And, in fact, to enjoin others to get it right just is in part to enjoin them to find out what getting it right means and requires in the relevant context. This emphasis motivates the shift towards what Cartwright et al (forthcoming), following the original suggestion of Eleonora Montuschi, call *Objectivity To Be Found*—it is part of injunctions to be objective to *find out* what being objective means and requires in a particular case. This helps us see what the notion of 'good judgement' means when we noted above that objectivity demands one make *good calls.* Good calls are calls that *get it right* relative to the needs of the context.

With this general picture on the table, let us turn to a more careful examination of what objectivity understood through the lens of *getting it right* requires and how, in particular, it can help us think through the duty of care in modelling.

4. Modelling Objectively

We follow Cartwright et al's (forthcoming) lead in identifying *two* general components demanded by objectivity. The first is to find the *right purposes*; the second, to find the right methods, or means, given those purposes.[11] Applied to modelling, we can think of the first as demanding that one answer the following question: what *should* this model be for? And, more broadly, the question: what should our *modelling* (or our *models,* plural) be for in this particular context? Finding the right methods, or means, on the other hand, will involve among other things: i) finding out what type or class of model should be used given the model's purposes, ii) finding out how certain features of a target system *should* be represented given the model's purposes and iii)

[9] Ballung stems from the German *Ballungsgebiet*: a congested urban area with ill-defined edges.
[10] Heather Douglas (2004, 454) identifies three operationally distinct types of objectivity and eight total 'senses' in which the concept is used, calling objectivity "an inherently complex concept, with no one meaning at its core."
[11] We use the term 'purposes' throughout, but we might just as well refer to ends or goals.

finding out how the various error-probabilities should be weighted given the model's purposes. So construed, a duty to model objectively is demanding; it requires earnestly seeking out both what it means and what is required to get it right in the case at hand. As we will see, sometimes this may require rejection of standard procedure and received wisdom.

Now, we claimed above that South Yorkshire Police failed to fulfil their duty of care at Hillsborough in part because of their failures to model objectively. We are now in a position to spell out further what this claim means. We suggest that South Yorkshire Police's failure to fulfil their duty of care was in part caused by their failure to find the *right purposes* for their modelling. How so? First, it seems clear that the police conceived of their role at the ground to be first and foremost a disciplinary one.[12] This was wrong. As the only party capable of ensuring crowd safety at the ground, the duty to do so naturally fell upon them.[13] Their failure to adequately focus on the full nature of their role thus led to their identifying the wrong purposes of their pre-match preparation; pre-match preparation which consequently concentrated on the narrow mechanics of crowd *control* rather than the much more complex task of crowd *safety.*[14] In sum, it was South Yorkshire Police's failure to find the right purposes for their pre-match preparation that constituted their failure to model objectively and this, in turn, helps explain their failure to fulfil their duty of care.

How does our emphasis on modelling objectively differ from other recent work on the duties involved in modelling? In a recent article focused on the responsibilities of modellers involved in epidemiological modelling, Eric Winsberg and Stephanie Harvard stress the importance of *clarifying a model's intended purposes.* They write:

> we argue that (1) clarifying the intended purpose of a model and (2) assessing its adequacy for that purpose are ongoing moral epistemic duties that must be upheld throughout the modelling process. (Winsberg and Harvard 2022, 512)

Winsberg and Harvard's proposal contributes to recent debates around the ethical responsibilities involved in Covid-19 modelling. Their emphasis upon a duty to *clarify* a particular model's intended purpose finds motivation in what they judge to be a persistent lack of clarity surrounding the purpose of Covid-19 models,[15] where they insist that such modellers must clarify whether a model is intended for prediction, forecasting or causal inference. As they rightly point out, clarification of such matters is crucial, for a failure to clarify the intended purpose of a model blocks one's ability to assess whether the model is good or bad; a model might be successful at prediction but useless for guiding interventions; we need to identify what we are aiming *at* in order to assess whether we have hit the target.

---

[12] In the 2000 case brought against Duckenfield and Murray, lead prosecutor Alun Jones, Q.C*.,* would claim that Murray and Duckenfield's "slow-motion negligence" was a product of their "mindset." That mindset, Jones effectively argued, was one that focused on the neutralisation of would-be troublemakers *over* the interests of crowd safety (Scraton 2004).

[13] See Taylor (1989, 36).

[14] We are making this claim on the basis of scholarship by Scraton (2004) but see also the witness statement by Superintendent Mole (South Yorkshire Police 1989). Here Mole describes the pre-match planning meeting undertaken by himself and Superintendent Duckenfield. The focus as described in this statement is exclusively on worries about supporter violence and does not mention issues pertinent to crowd safety such as overcrowding.

[15] For general discussion of these worries see Fuller (2021).

We think Winsberg and Harvard's contribution is important and commend their emphasis on what we believe they rightly call the 'moral-epistemic duties' in modelling. However, we stress two very important differences between our proposal and theirs. First, we wish to emphasise the importance of finding the *right* purposes rather than on clarifying *intended* purposes. As we will see in the section to follow, clarifying intended purposes is insufficient for fulfilling a modeller's duty of care. Second, we wish to depart from Winsberg and Harvard's analysis in emphasising that the duty to model objectively applies both to single models—*this* particular model right here—and integrated modelling efforts more broadly as they are carried out by entire institutions and organisations. The questions Winsberg and Harvard pose are questions targeted at specific models: is *this* particular model for prediction or causal interference? In contrast, we think of the duty of care to model objectively in much broader terms. The duty of care involves not only the earnestly seeking out of the right purposes for specific models but also the earnest seeking out of the right purposes for entire modelling efforts for which an individual model will be but one single part. The relevant injunction to *find the right purposes* is, on our view, an injunction that applies at both of these scales, or levels, simultaneously (see Fig. 1). We proceed in the section below with another case that will allow us to further bring out these two differences between our view and one that emphasises the clarification of intended purposes.

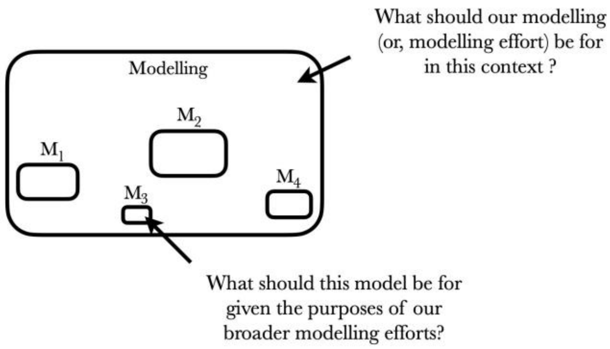

Fig. 1

5. Unobjective Modelling: The Vajont Dam Disaster

At 10:39pm on the 9th October 1963, inhabitants of the Italian town of Longarone were woken by the deafening sound of a landslide; 273 million cubic-meters of rock detaching from the northern side of Mount Toc. The rock would quickly find itself plunging into a 700-meter-high man-made reservoir at the rear of the monumental Vajont dam, sending a wall of water 20 meters high over the dam's edge. Over 2000 people living in Longarone and the surrounding villages would die. The dam was a magnificent feat of engineering, yet the landslide is one of the worst man-made catastrophes in history. With hindsight, the dam should never have been built the way that it was. So, what went wrong? How could such magnificent engineering go hand in hand with such catastrophic failure?

We focus below on two aspects on the modelling for the Vajont dam. First, we look at the initial feasibility studies conducted prior to the dam's construction. Second, we look at one instance of modelling that was carried out *after* the threat of landslides to the surrounding towns had been acknowledged. In the first case, we suggest that just as in the case of South Yorkshire Police at Hillsborough, the engineers for the Vajont dam were focusing on the *wrong purposes* for their initial feasibility studies; in this case, an *incomplete* conception of what this modelling was for. In the second case, we suggest that whilst the modellers were in this instance working with the right purposes, they had nonetheless failed to find the right means to achieve those purposes. Hence, in both cases we see a failure to model objectively.

5.1. Feasibility Studies for the Vajont Dam

The first feasibility studies began in the 1920s and ran until the land was finally purchased for development in 1943. In some ways, the Vajont valley at the bottom of Mount Toc was the ideal location for a hydroelectric dam. The acute angle of the gorge would make Vajont the tallest dam of its kind in the world; it would harness incoming water from three rivers, the Piave, the Mae and the Boite, and hold a total of 168 million cubic meters of water, supplying electricity for large portions of northern Italy. Yet, the location was not ideal; far from it. Residents living near the valley famously said as much. Indeed, the mountain was known by the locals for being geologically fragile and prone to landslides, with 'toc' (short for 'patoc') meaning *rotten* or *spoiled* in the local Italian dialect. Despite this information, the engineers at Società Adriatica di Elettricità (SADE), the company leading the project, did not demand that extensive geological studies of the region as a whole be carried out. Rather, their primary focus in these initial studies was the stability of the *abutment area*—the natural supporting structures either side of the would-be dam. Why were the investigations limited in this way?

Just as South Yorkshire Police were working with a misplaced emphasis on crowd control over crowd safety, so too, we suggest, were SADE working with the *wrong purposes* for these initial investigations. We here follow Pierluigi Barrotta and Eleonora Montuschi's (2018) account of the disaster in proposing that the engineers at SADE saw these investigations as focused, first and foremost, on the question: is this rock good enough to construct a dam that will *stand*? Yet, whilst obviously crucial, assessing the abutment area for threats against the stability of the dam was not the only purpose these initial studies should have served. The important questions to answer were *both:* i) is the rock good enough to construct a dam that will stand? And ii) can a reservoir of this scale be safely built here without threatening the geological stability of the region in a way that endangers human life? Whilst there can be no doubt that safety concerns must have occupied the SADE engineers during these initial feasibility studies, it nonetheless seems clear that the assessments of safety took a back seat relative to the task of assessing the ability for the abutment area to support a dam of this size. For instance, the chief engineer at SADE at the time, Carlo Semenza, is reported to have said:

> From a geological point of view the rocks [of the Veneto region] are generally very good [...]. Overall, limestone is honest because it reveals its flaws on its surface. (Carlo Semenza quoted in Gervasoni 1969, 11)

Here Semenza deployed a generic—that limestone reveals its flaws on its surface—in conjunction with his alleged observation that there were no major flaws exhibited by the rock to conclude that the region was sufficiently stable for construction of the dam.[16] Semenza's weak epistemic standards here only make sense, we suggest, if we understand the importance of safety assessments to be down-weighted relative to the importance of assessments of engineering feasibility.[17] Indeed, contrast Semenza's willingness to rely on the loose claim 'limestone wears its flaws on its face', with the magnificent levels of precision he demanded of the models used to design the dam itself—a system of 146 equations with as many unknown variables perfected and solved (Barrotta and Montuschi 2016).

Now, suppose that we are right and that Semenza and the other engineers at SADE were working with the wrong purposes when they carried out these initial studies, why is it not enough to say that they should have just *clarified the intended purposes* of their modelling? Why is it necessary to say that the duty of care involves a duty to find the *right* purposes? The problem with a focus on merely clarifying intended purposes is that it is entirely possible for a whole scientific community to clarify that the intended purpose of a particular modelling effort is X whilst the full, correct, purpose of that effort should, in actual fact, be Y. Indeed, it seems reasonable to assume that this was the very situation at SADE.[18] That is, it seems reasonable to assume that the intellectual climate at SADE was such that there might have been full agreement in a claim that the purpose of the feasibility studies was to ensure that the rock surrounding the abutment could support a dam of the relevant size. Yet, full agreement (and thus no lack of clarity) around the intended purposes of modelling is not sufficient for ensuring that modellers fulfil their duty of care. The duty of care requires more. Indeed, it demands that modellers go and actively find out what the model *should* be for, something that may require one to go *against* the consensus view of a particular community. Understanding the duty of care in modelling as a duty to model objectively brings out the crucial importance of *finding* and *seeking out.*

5.2. Scale Modelling to Assess Safe Reservoir Levels

So far we have said very little about the second aspect of modelling objectively finding the right methods (or means) to achieve the model's correct purpose. We turn to this now. As we will see,

[16] As Barrotta and Montuschi (2018) point out, this was to completely ignore the local knowledge attesting that this generic did *not* hold in the Veneto region.
[17] Barrotta and Montuschi write that in this initial stage, "the formidable engineering challenges of the project took precedence over the geological problems posed by the natural environment." (Barrotta and Montuschi 2018, 21)
[18] It seems like the same was likely also true at Hillsborough. That is, it is likely that South Yorkshire Police would have all *agreed* at the time that the purpose of their pre-match preparation was to facilitate effective crowd control on match day. Whilst there may not have been any lack of clarity about this purpose, however, it was nonetheless an *incorrect account* of the full set of purposes. This will be discussed below when we turn to the matter of how institutional failings may support unobjective modelling.

even in the case where one has identified the right purposes for a model a failure to find the right methods for achieving those purposes can lead to catastrophic failure. Making significant efforts to find the right purposes is therefore not sufficient for modelling objectively, one must also take pains to find the right tools for the job.

Signs of geological instability in the valley appeared shortly after construction of the dam began in 1957. On the 22nd of March 1959, three million cubic meters of material fell into the artificial lake of Pontesei, close to the Vajont Valley. The dislodged material was compact and the landslide swift; Semenza and his colleagues took notice. Semenza quickly commissioned a thorough geological survey of the region and the team's findings led them to a worrying hypothesis: there had been an ancient landslide on Monte Toc, and this landslide could potentially be reactivated by erosion from the dam's reservoir. Whilst such talk of major future landslides was met with initial pushback, by 1960 the evidence had become too great not to take the threat seriously. Indeed, by the end of 1960, a 2.5-kilometre-long fracture had opened on the northern slope of the Mount Toc. Shaped like an M, the fracture marked the contour of a huge sliding mass composed of two distinguishable bodies—bodies that would come to be known as the eastern and western lobes. The threat was now undeniable. Semenza and his colleagues turned their attention away from the question of *whether* major landslides were possible to the question of how catastrophe might be avoided.

One key question at this time was how high the reservoir could be safely filled, where 'safely filled' was effectively understood as 'filled so that there would be no danger to the surrounding towns of Longarone, Erto and Casso from landslides into the reservoir.' To answer this question, Semenza made an innovative move, requesting that Augusto Ghetti (then director of The Institute of Hydraulics at the University of Padua) build a 1:25 scale model to study the effects of landslides into the reservoir.[19] The purpose of the model was straightforward: find the level at which the reservoir could be safely filled. This was, we take it, the *right* purpose for this model. The problem, however, was that the model's design and the tests the Ghetti himself ran on the model were not sufficient to achieve this purpose. That is, whilst Ghetti's had the right purpose for his model, he did not take the right steps to achieve it.

There were two main issues with Ghetti's modelling. First, the materials used by Ghetti in the scale model[20] to simulate the real landslide mass could not provide the kinds of reliable inferences needed to make estimates of safe water levels in the reservoir. In particular, Ghetti used rounded gravel contained within metallic netting as a substitute for the real rockslide mass, a material that is significantly less compact and fell with significantly lower speeds than the real mass eventually would.[21] Second, Ghetti only ran simulations for *single* lobe failure scenarios despite the fact that

---

[19] The model was the first purpose-built scale model in Italy. (Barrotta and Montuschi 2018, 28)

[20] Whilst 'scale models' are so called because they are geometrically scaled versions of their target systems, the types of behaviour that scale models are used to investigate in science and engineering do not simply 'scale down' in the way geometric properties do. Using scale models to study the behaviour of materials or fluids demands sophisticated understanding of the transformations between the things happening in the model and the things being represented by those happenings. See Sterrett (2019) for an excellent treatment of the problems involved in scale modelling.

[21] Later modelling the at Vajont Dam was done to scale (i.e., 1:1)—indeed, it was done in the real setting, as the dam was filled, then drawn back as problems were observed, then refilled, three times. This to-scale modelling clearly did not deliver a good enough prediction either, due it seems to overoptimism both about the fill height that was safe and about the chance of massive landslides.

the real landslide would, in actual fact, involve the simultaneous failure of both lobes.[22] Altogether, Ghetti's modelling produced an estimate of 700 meters (above sea level) as the level at which there was to be no threat from overspill in the event of landslides. This would be the very level at which the dam was filled when the eventual landslide did occur, killing 2000.[23]

One might push back here and insist that whilst Ghetti's modelling turned out to have problems, Ghetti did nonetheless fulfil his duty of care to model objectively. For instance, one might argue that whilst relative to the actual facts his modelling *was* inappropriate, relative to what Ghetti could reasonably have taken those facts to *be,* his modelling was not inappropriate. Scale modelling was, after all, in its infancy when Ghetti designed his model, and simultaneous lobe failure scenarios were thought to be extremely unlikely at the time. We agree that the criticism of Ghetti's use of inappropriate proxy materials is difficult yet suggest that his failure to test simultaneous lobe failures *was* a failure of the duty of care given the stakes. The purpose of this model was to estimate levels at which the reservoir could be filled so that the residents of the surrounding towns would not be put at risk from landslides. Given that the M-shaped fracture and therefore the existence of the two independent bodies of material was well known, it was unacceptable not to test such joint failure scenarios. On this count, Ghetti thus failed to do what he should have done to achieve the model's purpose and this means that he failed to fulfil the duty of care to model objectively.

Now, whilst the case of Ghetti's scale modelling is one in which Ghetti had a duty of care to model objectively, in this as in most cases, the duty of care to model objectively will equally—if not more so—fall to institutions or collections of institutions. We naturally say it was 'SADE' who conducted the initial feasibility studies, for instance, and we naturally think of the failures in pre-match preparation at Hillsborough to fall upon 'South Yorkshire Police', where we understand this loosely to refer not only to the police officers themselves but also those institutions involved in hiring, training and overseeing these officers—all of these institutions were implicated in the failures in preparation that led to the disaster. We turn to discuss these matters of distributed responsibility next.

6. Objective Modelling: Whose Responsibility?

We have argued so far that when modelling is used to guide action that have effects in the real world, those involved in modelling have *a duty of care to model objectively*, where a duty to model objectively demands that one i) find out what the model (or modelling effort) should be for and ii) find out which tools and methods you should be using to achieve that purpose. In slogan form, objective modelling is modelling that *gets it right* and those involved in modelling have a duty of care to *earnestly try to get it right.*

[22] Here we have drawn on conclusions by Franci et al (2020) in their recent computer simulations of the Vajont disaster. Franci et al simulate Ghetti's own results and examine how variations in Ghetti's modelling techniques would have altered the safety estimate he provided. They identify the use of the gravel proxy and the focus on single lobe failure scenarios as the two key reasons why Ghetti's safety estimates proved inaccurate.

[23] Franci et al's (2020) simulation modelling suggests that Ghetti's figure of 700 meters was up to 75 meters *too high*; simultaneous lobe failure would have still reached over the top of the dam had the dam been filled to roughly 625 meters above sea level.

With this said, however, it is important to stress that getting it right does not happen in a vacuum. Because of this, we urge caution in affixing causal responsibility for failures to model objectively to particular individuals or even particular groups. For instance, it is very likely that other teams of engineers would have done exactly as Semenza and his colleagues did during their initial feasibility studies. As Barrotta and Montuschi (2018) note, a primary focus on the geological stability of the abutment area was, in fact, standard practice at the time—neither the norms within the engineering community nor the legal and regulatory norms in Italy required extensive geological investigations prior to the construction of dams. Hence, it is most likely that the other teams of engineers would have similarly adopted an incomplete set of purposes for their modelling, focusing, as Semenza and his colleagues did, on matters pertinent to the dam's construction over those pertinent to the safety of the inhabitants in the surrounding towns and villages. Similarly in the Hillsborough case. It is most likely that other police departments would have made similar mistakes in preparing for the FA semi-final that day.[24] When Superintendent Murray was asked why he and Duckenfield had not ordered the tunnel leading from Gate C to the pens to be closed, he replied that this "was something that did not occur to me at the time and I only wish it had" (Scraton 2004, 192). But we should think about *why* it did not occur to him or Duckenfield. What made it so natural for them to think and act as they did?

To begin to address these issues, we take a lead from UK child protection expert Eileen Munro in describing attempts to understand failures at child protection. Munro (2004) explains that when a tragedy occurs

> the standard response is to hold an inquiry, looking in detail at the case and trying to get a picture of the causal sequence of events that ended in the child's death…We are tracing a chain of events back in time to understand how it happened. (Munro 2004, 377)

What do we tend to conclude? Munro continues:

> The events that bring the investigation to a halt usually take the form of human error. (Munro 2004, 377)

Investigators look for deviations from the norms of professional behaviour, she notes:

> Practitioners did not comply with procedures or lapsed from accepted standards of good practice. (Munro 2004, 377-378)

This way of viewing failures is labelled the 'person-centered' approach. The trouble with the person-centered approach is that it isolates unsafe acts from their context, thus making it very hard to uncover and eliminate recurrent error traps within the system. By contrast, Munro urges a 'systems-centered' approach and argues that for children and young people protection is a *systems problem* (Munro 2004).

The US Institute of Medicine's *To Err Is Human: Building a Safer Health System* also urges a systems-centered approach, explaining:

[24] Concerns with hooliganism at the expense of crowd safety dominated 1980s football. Indeed, the very choice to design stadiums (like Hillsborough) with *enclosed* pens was a product of a mindset that neglected crowd safety in the interests of crowd control; the enclosed pens were designed to prevent pitch invasions but were, from the very beginning, a risk factor for lethal overcrowding (Scraton 2004).

> The title of this report encapsulates its purpose. Human beings, in all lines of work, make errors. Errors can be prevented by designing systems that make it hard for people to do the wrong thing and easy for people to do the right thing. Cars are designed so that drivers cannot start them while in reverse because that prevents accidents. Work schedules for pilots are designed so they don't fly too many consecutive hours without rest because alertness and performance are compromised. (Institute of Medicine 1999, ix)

Like Munro, the pamphlet urges:

> The focus must shift from blaming individuals for past errors to a focus on preventing future errors by designing safety into the system. (Institute of Medicine 1999, 5)

It is, we think, just these kinds of systems failure that we have seen in the cases of the Vajont Dam and Hillsborough and that we suggest are likely characteristic of failures of objective modelling more generally. Here 'the system' is constituted by the tangle of institutions and institutional norms into which modellers and modelling activities are embedded, from the institutions that train the modellers to those that fund research or oversee and regulate practice, including as well things like the cultural norms that pervade and the media pressure modelling may be consciously or unconsciously influenced by.

Systems failures that affect the objectivity of modelling will occur in various different institutions and at various different levels in different cases and they will surely take different forms and have different causes. But there is one fundamental factor that we think is likely to be widely relevant. This is due to what is often called 'silo-isation': the tendency of each area of expertise and practice to be closed in on itself and neglect what is going on elsewhere. Of course, sharing a common vision and language and common methods and practices is crucial to being able to get on with the hard work of getting a job done. But this can readily lead to a failure to take notice of what other areas of knowledge and practice and other perspectives have to offer, both in aid of getting the job in view done and in understanding just what the job is that should be done.[25]

Such failures to recognise relevant lessons—information, practices, methods, etc.—from elsewhere is often not deliberate nor even conscious. It is not that lessons from elsewhere are rejected or ignored but rather that often they do not come into view. Sometimes it is explicitly asserted that the practices within the silo are superior to what is available outside: *we* know how to do it. And sometimes the reverse: institutions explicitly undervalue what they have to offer (for instance when schools of qualitative and interpretive methodology agree that they can only *suggest* causal hypothesises but must leave confirming them to quantitative methods). But often this is not explicitly announced; it is just *presumed*, without notice. Altogether, we suggest that silo-isation thus often breeds a deep-seated failure of *intellectual humility* within the institutions that affect modelling, where this failure of intellectual humility involves an inappropriate presumption that the institution's own practices are the *right practices.*

We conjecture then that failures of intellectual humility in the institutions in which modelling practices are embedded can readily lead to failures of objective modelling. Whilst intellectual

[25] This was certainly visible in the ways in which the SADE engineers interacted with others with different forms of expertise. For instance, whilst Ghetti (a geologist) himself made mistakes in his own modelling, he nonetheless insisted that his safety level estimates were to be treated with extreme caution. This caution was *not* heeded by the SADE engineers who, disregarding Ghetti's own recommendations, took the 700-meter safety estimate at face value as a precise level to which they could safely fill the dam. See Barrotta and Montuschi (2018, 29).

humility has been primarily studied from the point of individual human agents,[26] we believe that many of the insights from this previous work fruitfully carry over to the study of institutions, and in particular, scientific institutions. Indeed, we think that many of the characteristic features of intellectual humility identified for the case of individual persons—e.g., an appropriate sensitivity to one's epistemic limitations and strengths; an awareness of one's own fallibility as a knower—can be rightly said to have been *absent* in the institutions involved in the modelling at Vajont and Hillsborough.

We end by noting that understanding the role of intellectual humility in scientific institutions is now an active research project both by ourselves[27] and others.[28] Two important goals of this research are to i) better spell out the causal pathways—including and along with silo-isation—through which a failure of intellectual humility might arise and ii) investigate potential strategies for fostering intellectual humility within scientific institutions. If we are right and failures to model objectively are often a result of a failure of intellectual humility in the institutions within which the relevant modelling practices are embedded, then understanding how to better foster intellectual humility is of critical importance.

7. Conclusion

When models are used to guide decisions that can possibly affect people's lives, those involved in modelling have a duty of care towards those who are likely to be affected by those decisions. In this paper we have spelled out some of what the duty of care in modelling requires. In particular, we have argued that modellers have a duty of care to *model objectively*. As we have seen in the cases of both the Vajont Dam and Hillsborough disasters, however, objective modelling does not occur in a vacuum. The activity of modelling is inherently embedded in a tangle of institutional norms, expectations, habits, patterns and practices, and modelling objectively will often if not always be a *systems-level problem.* We thus stress the importance of developing *systems-level solutions.* We have conjectured here that failures of *intellectual humility* in the institutions within which modelling practices are embedded is likely to be one such systems-level problem. If we are right in this, then future work must focus on developing appropriate solutions—ways of fostering intellectual humility within the institutions in which modelling practices are embedded.

[26] The last ten years have seen an increased interest in intellectual humility across philosophy, psychology, and the social sciences, especially encouraged by the John Templeton Foundation. In philosophy, intellectual humility has been proposed to consist in i) an awareness of one's fallibility as a knower (Spiegel 2012; Hazlett 2016), ii) a treatment of other intellects as worthy of as much respect as one's own (Priest 2017), and iii) an awareness of one's intellectual limitations in the right amount and for the right reasons (Church and Barrett 2016).

[27] In our John Templeton Foundation project, *The Successes and Failures of Science Through the Lens of Intellectual Humility: Perspectives from the History and Philosophy of Science* (PIs: Nancy Cartwright and Robin Hendry).

[28] Other current research on intellectual humility in scientific institutions includes work on the relationship between intellectual humility and the 'replicability crisis' in psychology as well as work on intellectual humility and the institutional biases against certain sub-fields in the social sciences such as those studying religion. See the projects by Mijke Rhemtulla, and Alexa Tullett, and Kimberly Rios, respectively at https://www.templeton.org/project/intellectual-humility. See also Higgins (2019) for recent work on intellectual humility in the institutions around social work and Davis et al. (2018) for discussion of intellectual humility in the context of institutional mechanisms such as peer review.

References

Atkinson, E. (2016, 26 April). "Hillsborough inquests: The role of police chief David Duckenfield." *BBC News Online.* https://www.bbc.com/news/uk-36141618. Accessed: 24th June 2022

Barrotta, P. and Montuschi, E. (2016). "Local Knowledge: Expertise, Method and Value." *PSA* 2016, Atlanta 2-6 November.

Barrotta, P. and Montuschi, E. (2018). "The dam project: Who are the experts? A philosophical lesson from the Vajont disaster." In Barrotta, P. and Scarafile, P. (Eds.), *Science and Democracy: Controversies and conflicts* (17-34), Amsterdam: John Benjamins.

Brown, M.J. (2019). "Is Science really value free? From Objectivity to Scientific Integrity." In McCain, K. and Kampourakis, K. (Eds), *What is Scientific Knowledge? An Introduction to Contemporary Epistemology of Science*, New York: Routledge.

Cartwright, N., Hardie, J., Montuschi, E., Soleiman, M. and Thresher, A. C. (Forthcoming). *The Tangle of Science*, Oxford: Oxford University Press.

Church, I. M. and Barrett, J. L. (2016). "Intellectual humility." In Worthington, Jr., E. L., Davis D.E. and Hook, J. N. (Eds.), *Handbook of Humility* (62-75), New York: Routledge.

Daston, L. and Galison, P. (2007). *Objectivity*. New York: Zone Books.

Davis, W. E., Giner-Sorolla, R., Lindsay, D. S., Lougheed, J. P., Makel, M. C., Meier, M. E., Matt E., Sun, J., Vaughn, L. A. and Zelenski, J. M. (2018). "Peer-review guidelines promoting replicability and transparency in psychological science." *Advances in Methods and Practices in Psychological Science* 1(4): 556-73

Douglas, H. (2000). "Inductive risk and values in science." *Philosophy of science* 67(4): 559-79.

Douglas, H. (2004). "The irreducible complexity of objectivity." *Synthese* 138(3): 453-73.

Franci, A., Cremonesi, M., Perego, U., Oñate, E. and Crosta, G. (2020). "3D simulation of Vajont disaster. Part 2: Multi-failure scenarios." *Engineering Geology* 279.

Fuller J. (2021). "What are the COVID-19 models modelling (philosophically speaking)?" *Hist Philos Life Sci* 43(47).

Gervasoni, A. (1969). *Il Vajont e le responsabilità dei manager,* Milan: Bramante Editrice.

Hacking, I. (2015). "Let's not talk about objectivity." In Tsou, J.Y., Richardson, A. and Padovani, F. (Eds.), *Objectivity in science* (19-33), New York: Springer.

Hazlett, A. (2016). "The Civic Virtues of Skepticism, Intellectual Humility, and Intellectual Criticism." In Baehr J. (Ed.), *Intellectual Virtues and Education: Essays in Applied Virtue Epistemology* (71-92), New York: Routledge.

Higgins, M. (2019). "The judgment in Re W (A child): national and international implications for contemporary child and family social work." *The British Journal of Social Work* 49(1): 44-58.

Institute of Medicine (1999). *To Err Is Human: Building a Safer Health System.* Washington, DC: The National Academies Press.

Munro, E. (2004). “Improving practice: child protection as a systems approach.” *Children and Youth Services Review* 27(4): 375-91.

Priest, M. (2017). “Intellectual Humility: An Interpersonal Theory.” *Ergo* 4(16): 463-480.

Rudner, R., (1953). “The scientist qua scientist makes value judgments.” *Philosophy of science:* 20(1): 1-6.

Scraton, P. (2004). “Death on the Terraces: The Contexts and Injustices of the 1989 Hillsborough Disaster.” *Soccer & Society* 5(2): 183-200.

South Yorkshire Police. (1989, 19 May). “Witness statement of Chief Superintendent Brian Mole.” *hillsborough.independent.gov.uk,* https://web.archive.org/web/20160706143904/http://hillsborough.independent.gov.uk/repository/docs/SYP000067080001.pdf. Accessed: 24th June 2022.

Spiegel, J. S. (2012). “Open-mindedness and intellectual humility.” *Theory and Research in Education* 10(1): 27-38.

Sterrett, S. G. (2019). “Scale Modeling.” In Michelfelder, D. P., and Doorn, N. (Eds.), *Routledge Handbook of Philosophy of Engineering* (394-407), New York: Routledge.

Taylor, P. (1989). *The Hillsborough Stadium Disaster, 15 April 1989: Inquiry by the Rt. Hon. Lord Justice Taylor, Interim Report.* London: The Stationery Office Books.

Winsberg, E. and Harvard, S. J. (2022). “Purposes and duties in scientific modelling.” *Epidemiol Community Health* 76: 512–17.